

# OPEN BIOLOGY

rsob.royalsocietypublishing.org

## Research



**Cite this article:** Bradshaw NJ *et al.* 2017 The *NDE1* genomic locus can affect treatment of psychiatric illness through gene expression changes related to microRNA-484. *Open Biol.* **7**: 170153.  
<http://dx.doi.org/10.1098/rsob.170153>

Received: 21 June 2017

Accepted: 9 October 2017

### Subject Area:

genetics/bioinformatics

### Keywords:

schizophrenia, gene expression, DISC1 network, NDE1, miR-484, pharmacogenetics

### Author for correspondence:

William Hennah

e-mail: [william.hennah@helsinki.fi](mailto:william.hennah@helsinki.fi)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3917938>.

THE ROYAL SOCIETY  
PUBLISHING

# The *NDE1* genomic locus can affect treatment of psychiatric illness through gene expression changes related to microRNA-484

Nicholas J. Bradshaw<sup>1,2</sup>, Liisa Ukkola-Vuoti<sup>3,5,6</sup>, Maiju Pankakoski<sup>3</sup>, Amanda B. Zheutlin<sup>8</sup>, Alfredo Ortega-Alonso<sup>3,5</sup>, Minna Torniaainen-Holm<sup>3,5</sup>, Vishal Sinha<sup>3,5,6</sup>, Sebastian Therman<sup>3</sup>, Tiina Paunio<sup>4,9</sup>, Jaana Suvisaari<sup>3</sup>, Jouko Lönnqvist<sup>3,9</sup>, Tyrone D. Cannon<sup>8</sup>, Jari Haukka<sup>3,7</sup> and William Hennah<sup>3,5,6</sup>

<sup>1</sup>Department of Neuropathology, Heinrich Heine University, 40225 Düsseldorf, Germany

<sup>2</sup>Department of Biotechnology, University of Rijeka, 51000 Rijeka, Croatia

<sup>3</sup>Mental Health Unit, and <sup>4</sup>Genomics and Biomarkers Unit, Department of Health, National Institute for Health and Welfare, 00271 Helsinki, Finland

<sup>5</sup>Institute for Molecular Medicine Finland FIMM, <sup>6</sup>Medicum, and <sup>7</sup>Department of Public Health, Clinicum, University of Helsinki, 00014 Helsinki, Finland

<sup>8</sup>Department of Psychology, Yale University, New Haven, CT 06520, USA

<sup>9</sup>Department of Psychiatry, University of Helsinki and Helsinki University Hospital, 00014 Helsinki, Finland

NJB, 0000-0001-5581-8828; WH, 0000-0002-4462-4208

Genetic studies of familial schizophrenia in Finland have observed significant associations with a group of biologically related genes, *DISC1*, *NDE1*, *NDEL1*, *PDE4B* and *PDE4D*, the 'DISC1 network'. Here, we use gene expression and psychoactive medication use data to study their biological consequences and potential treatment implications. Gene expression levels were determined in 64 individuals from 18 families, while prescription medication information has been collected over a 10-year period for 931 affected individuals. We demonstrate that the *NDE1* SNP rs2242549 associates with significant changes in gene expression for 2908 probes (2542 genes), of which 794 probes (719 genes) were replicable. A significant number of the genes altered were predicted targets of microRNA-484 ( $p = 3.0 \times 10^{-8}$ ), located on a non-coding exon of *NDE1*. Variants within the *NDE1* locus also displayed significant genotype by gender interaction to early cessation of psychoactive medications metabolized by CYP2C19. Furthermore, we demonstrate that miR-484 can affect the expression of CYP2C19 in a cell culture system. Thus, variation at the *NDE1* locus may alter risk of mental illness, in part through modification of miR-484, and such modification alters treatment response to specific psychoactive medications, leading to the potential for use of this locus in targeting treatment.

## 1. Introduction

The identification of genes that predispose to complex psychiatric traits is an important aspect in studying these conditions, however it is vital that this information is then used to improve our biological understanding and ultimately the treatment procedures for the disorders. This can be achieved through genetic studies in which, instead of using an end state diagnosis, alternative traits are employed that can measure a biological or pharmacological aspect of the condition.

Polygenic disorders, such as schizophrenia, are influenced by numerous interacting genetic factors, therefore identification of one candidate gene may aid in identification of others. This approach has been used in a large Finnish family

cohort, in which *DISC1* (disrupted in schizophrenia 1) was previously associated with schizophrenia [1,2], and which led to observation of association with four other genes (*NDE1*, *NDEL1*, *PDE4B* and *PDE4D*) [3,4] that encode protein binding partners of the *DISC1* protein [5–8]. The idea that such protein interaction partners of *DISC1* are encoded for by genes which show genetic interaction in mental illness is termed the *DISC1* network hypothesis. Specifically, multiple associations for psychiatric [2,9–11] and related endophenotypes, including memory [12], cognitive and neuroimaging [13] phenotypes, have been reported for *DISC1* in Finnish cohorts. By conditioning genome-wide linkage data for schizophrenia on *DISC1*, a peak of linkage at chromosome 16p was observed [3], near to *NDE1* (nuclear distribution element 1). This was followed up through association analysis at the *NDE1* locus, leading to the observation that a haplotype and its constituent SNPs associate with schizophrenia in this cohort, in a gender-dependent manner [3]. Genetic association for schizophrenia was, therefore, tested for other *DISC1* binding partners in this family cohort [4]. Although SNPs and haplotypes from six other genes were initially observed to associate, only variants in *NDEL1* (*NDE*-like 1, a close paralogue of *NDE1*) and in the phosphodiesterases *PDE4B* and *PDE4D* replicated when tested in a second, distinct sample from the cohort [4]. Recently, through further investigation of the roles of these variants in the *DISC1* network, the *NDE1* locus has been identified to increase risk of schizophrenia in this Finnish family cohort through interaction with high birth weight, a promising proxy measure for multiple pre- and/or perinatal environments [14].

The role of the *DISC1* network as a source of genetic risk for neuropsychiatric disorders is controversial due to the absence, to date, of evidence for their involvement in population based genomic studies of common variation [15,16]. However, these genes have been implicated at least within specific populations through strong evidence emerging from family based approaches and the studies of rare variants. In addition to the evidence from the Finnish family cohort, the *DISC1* and *PDE4B* genes are disrupted by chromosomal aberrations in Scottish families with major mental illness [8,17,18]. Furthermore, *NDE1* is independently implicated in major mental illness through its presence at 16p13.11, which is subject to duplications in schizophrenia [19–22], as well as being directly implicated through rare SNPs in patients [23]. The importance of the *NDE1* protein for neurodevelopment more generally has been dramatically demonstrated in individuals with biallelic loss of the functional *NDE1* gene, leading to severe microcephaly phenotypes, sometimes described in conjunction with lissencephaly or hydrocephaly [24–27]. Deletion of only one copy of the 16p13.11 locus, meanwhile, has been associated with neurological conditions including autism and epilepsy [28]. Recently it was shown that expression of mature miR-484, a microRNA that is encoded on an untranslated exon of *NDE1*, led to alterations in neural progenitor proliferation and differentiation, as well as behavioural changes in mice, thus implicating the microRNA in the phenotypes associated with 16p13.11 duplication [29].

We have previously studied the effect of *DISC1* network genetic variation on gene expression in a small ( $n = 30$ ) publicly available population cohort of the CEU (Utah residents with North and Western European ancestry) individuals. Using gene expression measured on the same individuals, by different laboratories, we were able to observe 528 genes as being replicable in their differential expression across 24

variants studied, of which 35 genes had pre-existing supporting evidence for a role in psychosis [30]. Intriguingly, seven of these affected genes were noted to be targets for drugs prescribed for psychiatric illness, leading to the hypothesis that these *DISC1* network variants, through their action on gene expression, may alter treatment outcome for medications designed to target these genes [30].

Here, in order to advance our understanding of the role these genes play in the aetiology of schizophrenia in Finland, we take this approach further. This is accomplished by using data on gene expression levels in case families in which these *DISC1* network genetic variants have been previously demonstrated to associate with schizophrenia [1–4], as well as by using data collected on how different psychoactive medications are used by the affected individuals within these families.

## 2. Material and methods

### 2.1. Study samples

The principal samples used here are part of a larger study of familial schizophrenia. These are Finnish patients born between 1940 and 1976, who were identified through the hospital discharge, disability pension and free medication registers [1,31]. The cohort totals 458 families (498 nuclear families) that contain 2756 individuals, of whom 2059 have been previously genotyped for the *DISC1* network genes [1–4]. Of these genotyped individuals, 931 are classified as affected with major mental illnesses using criteria from the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) [32]. These include 635 diagnosed with schizophrenia, 125 with schizoaffective disorder, 95 with schizophrenia spectrum diagnoses and 76 with other mental illnesses, including bipolar disorder and major depression. Here, two sub-sets of this larger familial sample were used as discovery cohorts, firstly a sample of 18 families (64 individuals, 23 broadly affected: 15 with schizophrenia, three with schizoaffective disorder, three with schizophrenia spectrum diagnoses, and two with other mental illnesses; 41 unaffected family members) for whom RNA was available to study gene expression level changes, and secondly all 931 broadly affected individuals were used to study medication use.

In order to replicate the gene expression results obtained from this family data, two independent cohorts were used. The first of these replication cohorts was a Finnish discordant twin pair sample ascertained for schizophrenia ( $N = 73$ ), for which information about recruitment and clinical evaluation has been described previously [33]. Briefly, the participants are 18 schizophrenia patients, their 18 unaffected co-twins and 37 control twins who have provided blood samples for gene expression analysis ( $N = 73$ ). The second replication cohort was the Genotype-Tissue Expression (GTEx) database ( $N = 338$ ), a publicly available resource for exploring the correlation between genotypes and gene expression across multiple tissues and in a genome-wide manner (accessed on 2 September 2016, [www.gtexportal.org/home](http://www.gtexportal.org/home)) [34]. To best match the source of the RNA used in the discovery cohort studies, data from whole blood were used for the GTEx tests.

### 2.2. Gene expression data

Total RNA was extracted from fresh blood samples from 82 individuals, with 18 individuals excluded from further analysis as their samples' RNA integrity number was lower than 8.

These individuals are from 18 families who were re-approached to provide RNA for gene expression analysis based on prior genetic observations in these families including *DISC1* [1], *RELN* [35] and *TOP3B* [36]. Genome-wide gene expression measures were assayed for this discovery cohort using the Illumina HumanHT-12 v4.0 Expression BeadChip. Of the 48 212 probes on the chip, 11 976 were significantly detectable at a threshold of  $p \leq 0.01$  in more than 90% of individuals. The expression data for these probes were processed using quantile normalization followed by  $\log_2$  transformation. Raw anonymous data regarding this family cohort can be accessed at the Gene Expression Omnibus (GEO) database (GSE48072). For the replication twin cohort ( $N = 73$ ), genome-wide gene expression data have been measured using the Illumina Human WG6 v3.0 chip, as reported previously in detail [37]. After quality control and data processing, identical to that used on the family data, 18 559 probes from this chip were significantly detectable.

### 2.3. Genotyping

In the discovery sub-cohorts used here, both genotype and expression data were available from 39 individuals, while 931 individuals had both genotype and medication data available. Thus, in order to ensure sufficient numbers of individuals for statistical testing, we only studied genetic variants that met specific minor allele homozygote frequencies in these sub-cohorts. In the discovery cohort for gene expression, a cut-off value for the minor allele homozygote frequency of  $\geq 10\%$  was implemented, providing five variants (*DISC1*: HEP3 haplotype (comprising SNPs rs751229 and rs3738401) and rs821616; *NDE1*: rs4781678, rs2242549 and rs1050162) with which to perform the analysis. For the discovery cohort for medication use, the frequency of the genetic variants was restricted to those with a minor allele homozygote frequency of at least 5%. This allowed seven variants from three *DISC1* network genes to be studied (*DISC1*: rs821616; *NDE1*: rs4781678, rs2242549, rs881803, rs2075512, and the haplotype of these four SNPs 'NDE1 Tag haplotype'; *PDE4B*: rs7412571).

The genotypes for the replication cohort of discordant twins were produced with the same method and at the same time as those described previously [1,3,4], with only two variants analysed (*NDE1*: rs2242549 and rs1050162) using the gene expression data. The analysis using the GTEx database as a replication cohort was conducted for all variants studied in the families, except for the *DISC1* haplotype.

All genetic variants analysed in this study have been implicated by previously described evidence as being associated with the aetiology of schizophrenia in this cohort [1–4], with variants in both *DISC1* and *NDE1* having displayed prior gender-dependent effects [1,3,12]. Therefore, no multiple test correction has been applied to correct for the multiple testing across variants or gender interaction models, as they can all be considered hypothesis based. However, because we are screening alternative phenotypes in a hypothesis-free manner, we have applied the measures described in the following sections in order to further characterize these *a priori* variants.

### 2.4. Association between genome-wide expression levels and genotypes

For the discovery cohort used to study gene expression, a mixed effect linear regression model was fitted for each

probe and genotype using R (RStudio version 0.99.489) lme4 package [38], after correcting for gender, age, diagnostic status and family or twin status effects as a covariate. This analysis was performed separately but identically for the discovery cohort and the replication twin cohort. We used the false discovery rate (FDR) method in place of a family-wise error rate (FWER). FDR is widely applied for microarray analyses because it allows more genes to be extracted for further exploration, and was performed using the qvalue package in R [39] to estimate the FDRs of  $q \leq 0.05$ . The *post hoc* power of our small familial discovery cohort to detect gene expression changes was estimated using R package ssize.fdr [40]. The GTEx database was mined using its own in-built test procedure, entering in a list of gene IDs to be tested against our SNPs of interest. Data from whole blood was used in order to replicate only those genes identified as significantly altered in their gene expression levels at a cut-off of  $p \leq 0.05$ . When testing for replication, probes significant ( $p \leq 0.05$ ) in the discovery cohort prior to application of FDR were studied. See figure 1 for a flow chart of analysis.

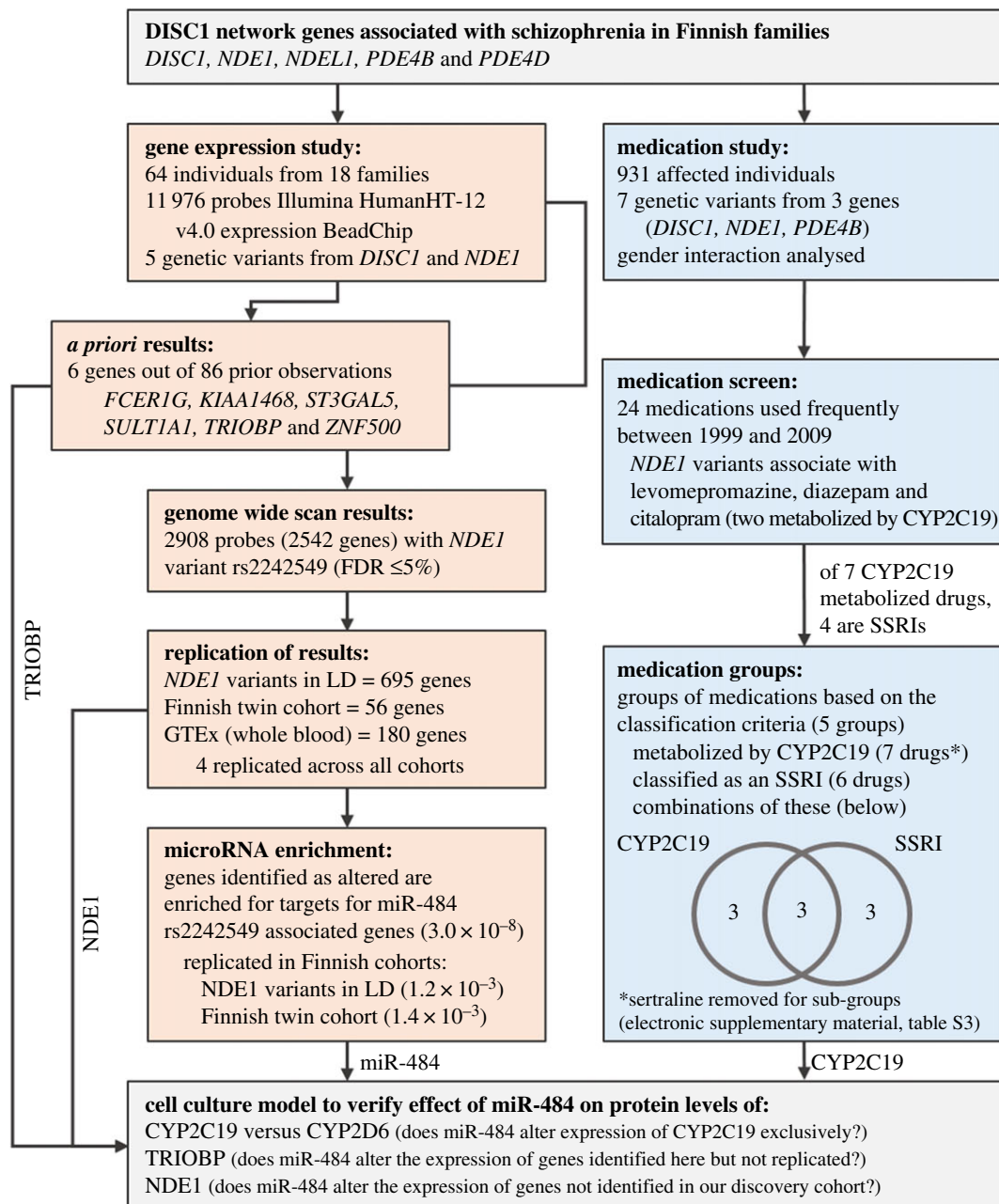
### 2.5. MicroRNA target prediction and enrichment analysis

A comprehensive list of predicted 3'UTR targets for miR-484 were obtained from the miRWalk database [41], considering only genes predicted by at least six of the 12 programs. Ingenuity pathway analysis (QIAGEN Redwood City, [www.qiagen.com/ingenuity](http://www.qiagen.com/ingenuity)) was used to analyse potential enrichment of these genes among those significantly altered in their expression levels by *NDE1* SNPs in the discovery and replication cohorts.

### 2.6. Medication data

Prescription medication data from the 931 affected individuals of the schizophrenia family cohort for the period 1 January 1996 to 31 December 2005 was obtained from the Finnish National Prescription Register of the National Social Insurance Institution (SII) [42], thus forming the discovery cohort for medication use. In this cohort 'medication use' is based on purchases of prescribed psychoactive drugs for which the SII have paid a reimbursement. Data from this register contain information on date of purchase and the dose, stated as the international standard daily defined dose. Thus, using this data, medication periods were defined according to method 4 proposed by Mantel-Teeuwisse *et al.* [43], multiplying defined daily dose by a factor of 1.1 and filling 15-day gaps between medication periods. This medication period data was used to determine the probability of cessation of each drug by genotype and converted into a binary variable using three months as a cut-off. This three months cut-off reflects that an individual either purchased more of the same medication, or purchased a different medication within a three-month period after the original purchase date. Only psychoactive medications with at least 15 instances of use for three months or less were taken for association analysis of the individual drugs, meaning that, of the possible 101 psychoactive drugs for which data was available, only 24 were analysed. When data from multiple drugs were combined in order to study classes of medication, all medications for which data are available were used, regardless of individual frequency. In order to account for the fact that





**Figure 1.** Schematic flow chart of the analyses undertaken in this study.

some medication periods may come from the same individual, analysis of the medication usage used logistic regression with GEE-estimation (generalized estimating equations), as used by the geepack package for R [44,45]. Bonferroni correction was used to correct for the multiple tests in the analysis of the individual medications (24 tests) and the groups of medications (five tests). While  $p$ -values are presented unadjusted, only those that would surpass the Bonferroni correction-based thresholds ( $p \leq 0.0021$  and  $p \leq 0.01$ , respectively) are highlighted. See figure 1 for a flow chart of analysis.

## 2.7. Cell culture and western blotting

To determine whether miR-484 had the potential to affect the expression of selected proteins in a human cell-based system, NLF neuroblastoma cells (Children's Hospital of Philadelphia) were grown in RPMI 1640/10% fetal calf serum/2 mM L-glutamine (all from Thermo Fischer Scientific) and transfected with 50 nM of either a mimic of mature miR-484 (QIAGEN, Sy-hsa-miR-484) or a negative control microRNA (QIAGEN,

AllStars Negative Control microRNA) using Lipofectamine 2000 (Thermo Fischer Scientific) according to the manufacturer's instructions. After 48 h, cells were lysed using PBS/1% Triton X-100/20 mM MgCl<sub>2</sub> containing protease inhibitor cocktails and DNaseI. Lysates were western blotted and proteins detected using the following antibodies: anti- $\alpha$ -actin (Sigma, A2066), anti-CYP2C19 (Novus Biologicals, NBP1-19698), anti-CYP2D6 (Abnova, H00001565-B01P), anti-NDE1 (Proteintech, 10233-1-AP), anti-TRIOBP (Sigma, HPA019769) and anti- $\alpha$ -tubulin (Sigma, T9026). IRDye secondary antibodies were used (LI-COR) and the signal visualized using an Odyssey CLx infrared imaging system (LI-COR). Protein bands of interest were then isolated and quantified using IMAGE STUDIO software (LI-COR). In the case of TRIOBP and CYP2D6, which are each encoded for by multiple splice variants [46–48], the species at 72 and 56 kDa, respectively, were selected for analysis as they are known to be biologically active [49,50]. In the case of NDE1, all of the bands between 40 and 45 kDa, which are consistent with its large number of known phosphorylation events [51], were measured and pooled. All membranes were probed

with secondary antibodies alone first to ensure specificity of signal. Antibody signals were normalized to actin as a loading control. Mean fold-changes between the control and miR-484-treated samples were calculated from seven to eight technical replicates, performed from the same batch of cells. Three independent experiments of this type were then performed for biological replication, and the results compared by two-tailed paired Student's *t*-test.

### 3. Results

#### 3.1. Replication of gene expression changes from previous studies

In our previous analysis of gene expression in association with DISC1 network variants using publicly available data on the CEU (Utah residents with North and Western European ancestry) individuals, 86 genes were found to be differentially expressed across the three DISC1 network variants that could be directly replicated here (*DISC1* rs821616, 18 genes; *NDE1* rs4781678, 39 genes; *NDE1* rs2242549, 29 genes) [30]. To verify these results, made in a cohort of 30 individuals using technical replications, these three variants were tested again, this time using the Finnish family cohort. In total, six of the gene expression changes previously reported could be replicated, for the genes *TRIOBP*, *ZNF500*, *KIAA1468*, *FCER1G*, *SULT1A1* and two probes for the *ST3GAL5* gene. These were all in association with the status of the *NDE1* gene locus (rs2242549). In the previous analysis, *NDE1* rs2242549 associated with changes in gene expression levels of 29 genes, thus these six genes represent a 21% overlap.

#### 3.2. Differentially expressed genes associated with DISC1 pathway genotype

To further investigate the effect of DISC1 network variants on gene expression, we used this Finnish family cohort as a discovery sample to investigate the association of previously positive *DISC1* and *NDE1* variants with genome-wide gene expression of 11 976 probes. Notably, the *NDE1* SNP rs2242549 was significantly associated with gene expression levels of a large number of the probes (table 1 and electronic supplementary material, S1, figure S1). Specifically, 3824 probes representing 3314 distinct genes showed uncorrected association with the *NDE1* SNP rs2242549 ( $p \leq 0.05$ ), of which 2908 probes, representing 2542 distinct genes, were associated at FDR  $\leq 5\%$ . We also verified that, despite the size of the discovery cohort, it had sufficient power to detect the differentially expressed genes associated with the rs2242549 variant (electronic supplementary material, figure S2). By contrast, no probes were significantly altered in expression levels in association at FDR  $\leq 5\%$  with either the *DISC1* variants tested (SNP rs821616 of the 'HEP3' haplotype comprising rs751229 and rs3738401) or with the other *NDE1* SNPs tested (rs4781678 or rs1050162).

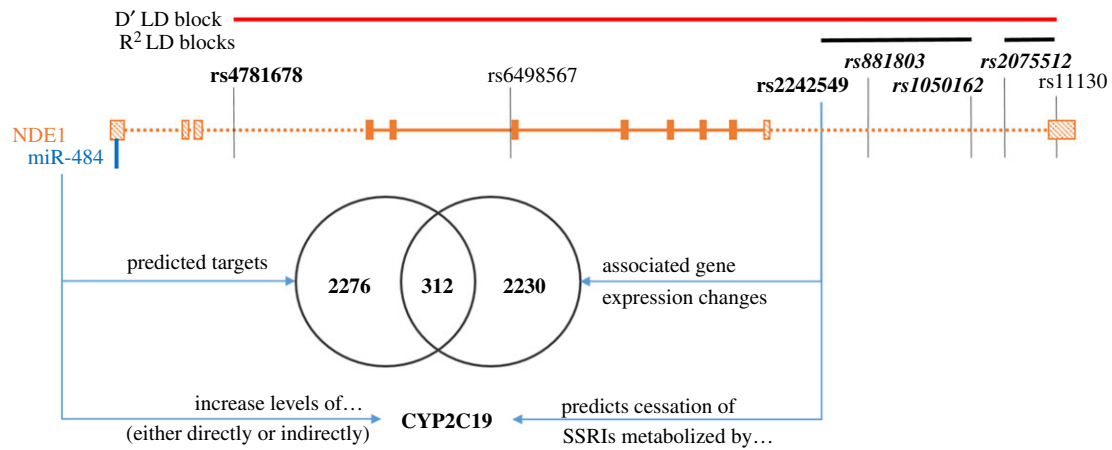
#### 3.3. Replication of gene expression changes associated with NDE1 genotypes

In order to replicate these observations, we pursued three lines of supporting evidence. Firstly, we noted that although all variants tested are in high D' linkage disequilibrium (LD) and thus

**Table 1.** Number of probes and genes significantly altered by variants in the DISC1 network, how they replicate across different cohorts, and how they overlap with the predicted targets of miR-484. n.a., value was not applicable as it was not tested; n.s., value was not significant nor returned by IPA.

gene	variant	observations		gene replications				replication cohorts	
		discovery cohort		discovery cohort		a priori observations		Finnish SCZ twins	
		probe $p \leq 0.05$	gene $p \leq 0.05$	probe $q \leq 0.05$	gene $q \leq 0.05$	variants in high LD	Hennah & Porteous [30]	n.a.	GTEx (whole blood)
<i>DISC1</i>	haplotype HEP3	432	419	0	0	n.a.	n.a.	n.a.	n.a.
	rs821616	1332	1265	0	0	n.a.	1	n.a.	49
	rs4781678	340	327	0	0	n.a.	1	n.a.	18
<i>NDE1</i>	rs2242549	3824	3314	2908	2542	752 (695 genes)	7 (6 genes)	56	180
	rs1050162	1071	985	0	0	752 (695 genes)	n.a.	13	60
enrichment for predicted targets of microRNA-484 <sup>a</sup>									
rs2242549			$2.2 \times 10^{-13}$		$3.0 \times 10^{-8}$	$1.2 \times 10^{-3}$		$1.4 \times 10^{-3}$	$5.5 \times 10^{-2}$
rs1050162			$5.6 \times 10^{-4}$			$1.2 \times 10^{-3}$		n.s.	$1.5 \times 10^{-1}$

<sup>a</sup>Targets predicted by at least 6 out of 12 prediction tools summarized by miRWalk were uploaded to ingenuity pathway analysis (IPA), thus enabling a test for enrichment when the subsequent probe/gene lists were studied.



**Figure 2.** Schematic of the *NDE1*/miR-484 locus, and summary of findings from this study. The exonic structure of the *NDE1* gene and the relative location of miR-484, both on the + strand of chromosome 16p13.11. For *NDE1*, exons present in all transcripts encoding full-length (300 amino acid or longer) *NDE1* protein are filled in with solid colour, while optionally transcribed exons are indicated with a striped pattern. The SNPs genotyped in the large Finnish family cohort are labelled, with the LD block structure summarized (detail in electronic supplementary material, figure S3). SNPs analysed in this study are labelled in bold type, with those SNPs that could only be studied in one of the sub-cohorts being further indicated with italics.

tag the underlying haplotype structure, only one SNP was in high  $R^2$  LD with rs2242549 (rs1050162,  $R^2 = 0.88$ ,  $n = 1891$  individuals genotyped at both loci; electronic supplementary material, figure S3, as determined using Haploview [52,53]) and could, therefore, be assumed to act as an internal replication of observations. Thus, within the discovery cohort, of the genes whose expression levels were associated with rs2242549, 752 probes representing 695 genes were also significantly associated with rs1050162 prior to application of the FDR ( $p \leq 0.05$ ). Secondly, we used existing data from a twin cohort for schizophrenia from Finland [33,37] as a replication cohort, identifying 56 probes, each representing a different gene, with replicable significant alteration in their expression levels associated with rs2242549 (table 1 and electronic supplementary material, S1). Finally, we used the publicly available database GTEx as an additional independent replication cohort, from which we were able to directly test 2651 of the 3314 genes identified (those probes with an official gene name), confirming that 180 genes also display significant gene expression changes related to the *NDE1* SNP rs2242549 in this database (table 1 and electronic supplementary material, S1).

In total, 794 probes representing 719 genes had supporting evidence from at least one additional source for their changes in expression related to the *NDE1* variant rs2242549, of which 76 probes from 73 genes had supporting evidence from more than one source, and four probes from independent genes (*ITGB5*, *OVGP1*, *PGRMC1*, *TST*) had supporting evidence from all three sources, that is SNPs in high LD in the discovery cohort, as well as independent replication cohorts of Finnish twins and the GTEx database.

### 3.4. Enrichment of miR-484 target genes among genes whose expression is associated with *NDE1* genotype

The finding that expression levels of such a large number of genes could be altered by a single genetic locus was surprising, especially given that the principal functions of the *NDE1* protein are not known to be in gene regulation. The *NDE1* locus also encodes for a microRNA (miR-484), which is located on a non-coding 5' exon of the longest splice variant of the

*NDE1* gene (figure 2). Since the major function of microRNAs is in the regulation of expression of other genes it is the most likely explanation for the sheer number of expression changes observed to associate with these SNPs. We therefore investigated whether the set of genes whose expression is altered by these *NDE1* SNPs overlapped with those genes predicted to be targets of miR-484.

Using the miRWalk database [41], 16 027 gene targets are predicted for miR-484, of which 2588 are predicted by at least six or more of the 12 independent prediction programs used in the database (data collated June 2016). Upon examining the list of genes whose expression is altered by SNPs in the *NDE1* locus at FDR  $\leq 5\%$  in our discovery cohort, these probes were indeed seen to be enriched for predicted targets of miR-484 ( $p = 3 \times 10^{-8}$ ). Employing the same tests to the three replication studies described above, the enrichment in miR-484 targets was also present for the set of these genes whose expression level is significantly associated with the genotype of both *NDE1* SNPs rs2242549 and rs1050162 (which were in high LD), and among those genes which could also be observed in the replication twin cohort. In contrast, they were not enriched among those genes replicated by data from the GTEx database (table 1).

### 3.5. Medication cessation associated with *NDE1* genotype

In our previously published analysis we found *NDE1* genotypes that significantly associated with expression levels of targets for particular medications relevant to mental health [30]. We therefore tested whether *NDE1* genotypes were also associated with cessation of specific medications in a discovery cohort consisting of all affected individuals from the Finnish family cohort. Screening all of the medications frequently used within this discovery cohort, we observed an association between *NDE1* rs4781678 genotype and early cessation of use of the antipsychotic levomepromazine (OR = 4.13 per C allele; 95% CI = 1.72–9.91;  $p = 0.00090$ ). When analysed in interaction with gender, association was further observed across *NDE1* genotypes with early cessation of the use of diazepam and citalopram (electronic supplementary material, table S2),



two drugs that share a common principal metabolizing enzyme, CYP2C19 (cytochrome P450 2C19, electronic supplementary material, table S3) [54–57]. We therefore asked whether *NDE1* genotype was associated with the subset of medications metabolized by the CYP2C19 enzyme. Since four out of seven of the drugs metabolized by CYP2C19 were selective serotonin reuptake inhibitors (SSRIs), we studied these as a separate group as well as further separated based on CYP2C19 metabolism. No significant interaction was observed for the grouping of all drugs metabolized by CYP2C19, however a genotype by gender interaction was noted when all SSRIs were grouped together (rs2075512, OR = 0.37; 95% CI = 0.17–0.79;  $p = 0.010$ ). When all SSRIs metabolized by CYP2C19 were tested, this genotype by gender interaction became significant (ranging from: OR = 0.27 to 0.31; 95% CI = 0.11 to 0.13–0.64 to 0.71;  $p = 0.0030$  to 0.0060) for four out of the five *NDE1* markers tested, while no interaction was noted for SSRIs not metabolized by CYP2C19 (table 2 and electronic supplementary material, figure S4). We analysed the remaining drugs metabolized by CYP2C19 together as another grouping (non-SSRIs metabolized by CYP2C19). Interestingly, a significant interaction was observed (ranging from: OR = 3.33 to 5.82; 95% CI = 1.44 to 2.25–7.25 to 15.0;  $p = 0.0013$  to 0.00030) for all five *NDE1* markers (table 2 and electronic supplementary material, figure S4). In this case, however, the gender effect is reversed, with SNPs being associated with cessation among females, in contrast to SSRIs metabolized by CYP2C19, for which SNPs were associated with cessation among males.

### 3.6. The effect of miR-484 on CYP2C19 in cultured cells

Given that a major effect of the *NDE1* locus variants examined here seems to be to alter the expression of genes targeted by miR-484, presumably due to altered expression of this miR-484, we hypothesized that the pharmacological consequences of these variants were also likely to occur through miR-484. For this to be the case, miR-484 would need to affect the levels of CYP2C19 protein expression, and thus be able to alter its metabolic activity effects on psychoactive medication.

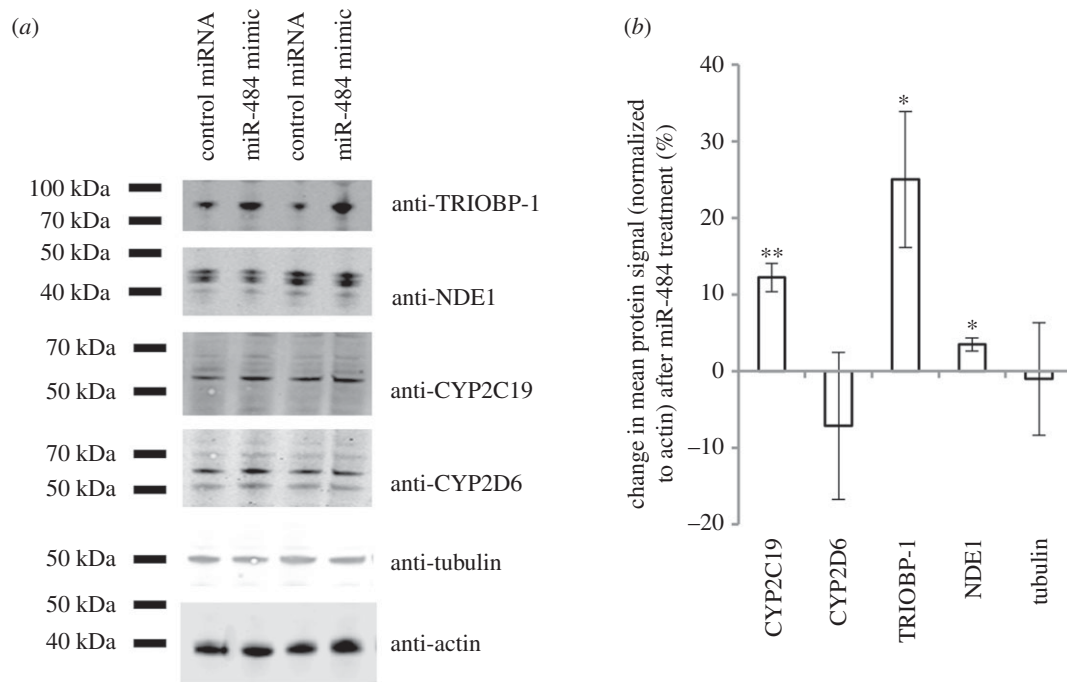
We therefore conducted a proof-of-principle experiment in NLF human neuroblastoma cells, into which we transfected a mimic of the mature form of human miR-484. Protein levels of CYP2C19 were significantly upregulated following miR-484 transfection, when compared with transfection with a negative control microRNA (figure 3). In contrast, no effect on the expression of another principal metabolizing enzyme of psychoactive medications, CYP2D6, was seen, indicating that this is a specific effect. That the levels of these proteins are altered, in a manner that closely matches to the pharmacogenetic data, strongly implies that miR-484 influences the expression of CYP2C19, but not CYP2D6, transcripts, although additional data at the transcript level would be required to confirm this.

Finally, we also used the same system to investigate two proteins that are predicted targets of miR-484, each of which illustrates a potential false negative in this study, possibly as a result of power limitations. The first of these, *TRIOBP*, was observed to be significant in our discovery cohort but not in the replication cohorts, while the other is *NDE1* itself, which was highly significant in the larger GTEx database ( $\beta = 0.24$ ; s.e. = 0.037;  $t$ -statistic = 6.6;  $p = 2.3 \times 10^{-10}$ ), but not originally observed in our discovery cohort. *TRIOBP* was further selected as it is also an example of the six genes whose

**Table 2.** Results for the association of *NDE1* variants with groups of medications based on their metabolism by the CYP2C19 enzyme and/or selective serotonin reuptake inhibitor (SSRI) class status, showing the  $p$ -values and odds ratios (and 95% CIs) for the interaction model.  $p$ -Values  $\leq 0.01$  are below the Bonferroni correction threshold for the five groups tested.  $p$ -Values and their respective ORs that are below the Bonferroni threshold are in bold. Medications included in each group analysis can be found in the electronic supplementary material, figure S4 and table S3.

<i>NDE1</i> variant	SNP $p$ -value	SNP $\times$ gender $p$ -value	OR (95% CI)
all psychoactive drugs metabolized by CYP2C19 ( $N = 510$ –581, no. of instances: 843–945)			
rs4781678	0.79	0.90	0.97 (0.55–1.69)
rs2242549	0.56	0.74	1.10 (0.64–1.86)
rs881803	0.29	0.15	1.53 (0.86–2.73)
rs2075512	0.64	0.79	0.93 (0.53–1.63)
haplotype	0.35	0.73	1.10 (0.63–1.93)
SSRIs ( $N = 357$ –402, no. of instances: 496–563)			
rs4781678	0.31	0.033	0.43 (0.20–0.93)
rs2242549	0.92	0.031	0.46 (0.23–0.93)
rs881803	0.41	0.29	0.67 (0.32–1.40)
rs2075512	0.97	<b>0.010</b>	<b>0.37 (0.17–0.79)</b>
haplotype	0.87	0.029	0.42 (0.19–0.91)
SSRIs metabolized by CYP2C19 ( $N = 282$ –318, no. of instances: 357–395)			
rs4781678	0.42	<b>0.003</b>	<b>0.27 (0.11–0.64)</b>
rs2242549	0.47	<b>0.006</b>	<b>0.30 (0.13–0.70)</b>
rs881803	0.51	0.064	0.44 (0.19–1.05)
rs2075512	0.39	<b>0.005</b>	<b>0.31 (0.13–0.70)</b>
haplotype	0.95	<b>0.006</b>	<b>0.30 (0.13–0.71)</b>
SSRIs not metabolized by CYP2C19 ( $N = 123$ –145, no. of instances: 142–169)			
rs4781678	0.48	0.71	1.36 (0.27–6.87)
rs2242549	0.30	0.58	1.37 (0.45–4.23)
rs881803	0.58	0.23	2.24 (0.60–8.38)
rs2075512	0.10	0.88	0.88 (0.19–4.13)
haplotype	0.66	0.77	1.25 (0.27–5.76)
non-SSRIs metabolized by CYP2C19 ( $N = 318$ –355, no. of instances: 411–459)			
rs4781678	0.24	<b>0.0051</b>	<b>3.33 (1.44–7.73)</b>
rs2242549	0.62	<b>0.0013</b>	<b>3.42 (1.61–7.25)</b>
rs881803	0.63	<b>0.00030</b>	<b>5.82 (2.25–15.0)</b>
rs2075512	0.92	<b>0.0064</b>	<b>3.63 (1.44–9.17)</b>
haplotype	0.22	<b>0.00030</b>	<b>4.93 (2.07–11.76)</b>

expression was associated with *NDE1* variation both here and in our previous analysis [30]. In both cases, protein levels were subtly, but statistically significantly, altered following treatment with the miR-484 mimic (figure 3), in comparison to the negative control microRNA.



**Figure 3.** The effect of a miR-484 mimic on protein expression in human NLF neuroblastoma cells. (a) Sample western blots showing levels of six proteins in the lysates of cells which had been transfected with either a mimic of miR-484 or with a negative control microRNA. (b) Quantification of three independent experiments, each comprising 7–8 internal replicates. All proteins were normalized to actin. \* $p < 0.05$ , \*\* $p < 0.01$ .

## 4. Discussion

Here we have demonstrated that variations within the *NDE1* locus, encoding a protein of the DISC1 network of protein interaction partners, can affect both gene expression levels and medication usage of psychoactive drugs used to treat major mental illnesses. Specifically, two SNPs in high LD are associated with replicable expression changes in a large number of genes, and with early cessation of psychoactive medications metabolized by CYP2C19 in a gender-dependent manner. We propose that these observations are linked through the involvement of miR-484. This microRNA is encoded for within the 5' untranslated exon of the longest splice variant of *NDE1*, and the one which is most abundantly expressed, at least in cell culture [58]. Notably, the list of genes whose expression changes are associated with these *NDE1* locus variants is significantly enriched for predicted targets of the microRNA, while expression of the CYP2C19 protein has been demonstrated *in vitro* to be significantly increased following addition of a mimic of mature human miR-484. This latter result was obtained from neuroblastoma cells. While such immortalized cells represent an atypical genetic environment, and some caution must be taken in directly applying them to *in vivo* situations, they represent a useful tool to perform proof-of-principle experiments in a reductionist manner. Nevertheless, the most promising explanation for the observations described here would, therefore, be that variation at the *NDE1* locus affects gene expression and medication metabolism in large part through effects of the variant on miR-484, and this may even be behind our prior observations at this locus of association to schizophrenia. It is interesting to note that a 1.45-fold increase in miR-484 has previously been reported in the superior temporal gyri of patients with schizophrenia [59]. Regarding the mechanism by which SNPs in the *NDE1* locus may affect expression of miR-484, it is noteworthy that miR-484 is transcribed from the same strand as *NDE1*, with both the

microRNA and its promoter [60] being found within a 5' untranslated exon of *NDE1*, which is included in some but not all *NDE1* splice variants. It is therefore very likely that any event which affects either the expression or alternate splicing of *NDE1* would also impact upon miR-484 expression. That many of the SNPs of the *NDE1* haplotypes are found in the 3'-most intron of *NDE1* (figure 2), which contains many motifs that regulate *NDE1* splicing [61], therefore allows speculation that the SNPs could affect miR-484 expression indirectly through modulation of *NDE1* splicing events.

The 16p13.11 locus, in which the *NDE1* gene is found, is prone to copy number variations (CNVs), with these 16p13.11 CNVs having been repeatedly associated with psychiatric and neurological disorders [19–22,28]. This locus contains multiple genes, however *NDE1* has been considered among the most promising candidates to be involved in these disorders due to its known critical role in neurodevelopment (reviewed: [62]). Therefore, our observations here, although of specific SNPs at the *NDE1* locus, highlight disruption of miR-484 as a potential functional consequence also of those CNVs. These results partially parallel recent findings that expression of mature miR-484 led to alterations in neural progenitor proliferation and differentiation, as well as behavioural changes in mice, thus implicating the microRNA in the phenotypes associated with 16p13.11 duplication [29]. While *NDE1* over-expression was not seen to have a gross effect on neuronal progenitor proliferation under similar circumstances, given the severe neurological consequences of biallelic disruption of the *NDE1*, but not miR-484, reading frame [24–27], there is still a potential role for *NDE1* in the conditions associated with 16p13.11 duplication. Additionally, relatively mild phenotypic effects would be needed to explain the fact that while associated with schizophrenia risk, most carriers of the CNV do not develop the condition [19–22]. Nevertheless, it can be speculated that the mechanism implicated here through SNP association, that miR-484 drives large gene expression



changes, could also be a consequence of the duplication of this locus.

This study initially sought to replicate our previous work on the effect of *DISC1* network variants on gene expression changes in the general population, using publicly available data on the CEU (Utah residents with North and Western European ancestry) individuals [30]. Of these previously identified 86 genes, we were able to replicate the observed changes in expression of six genes, including expression changes of two probes for *ST3GAL5* and a probe for *TRIOBP*, the only one of these genes to be a predicted target for miR-484. All of these were in association with the *NDE1* SNP rs2242549, and represented a 21% overlap with the findings from the prior study [30]. When we tested for gene expression alterations across the genome we identified a large number of probes ( $n = 2908$ ) representing 2542 genes whose expression levels associated with variants at the *NDE1* locus, specifically with the SNP rs2242549. A large proportion of these (752 out of 3824 probes, 695 out of 3314 genes) were significantly altered by another *NDE1* SNP (rs1050162), which is in high LD with rs2242549. Yet replication attempts in independent cohorts, although providing validation for some genes (56 in an independent Finnish schizophrenia cohort and 180 using the GTEx database), did not provide unilateral confirmatory evidence, with the exception of four genes (*ITGB5*, *OVGP1*, *PGRMC1*, *TST*) identified across all three datasets tested. This lack of replication, combined with our new observations and their biological relevance through miR-484 to the *NDE1* locus, suggests that, although the variants studied here are common to many populations, their relationship to potential functional mutations at this locus, and their specific biological consequences associated with schizophrenia and gene expression changes, may be unique to this Finnish family cohort [63]. This population difference may account for the lack of replication of most of the previously observed genes in the CEU population [30], and the lack of enrichment for miR-484 targets in the GTEx database. This is consistent with *DISC1* variation playing a genetically heterogeneous role in the general incidence of schizophrenia, lacking common illness-associated variations which could be detected by genome-wide association studies [15,16] of global populations, but providing strong evidence for a role in the condition within specific populations and family studies [63].

Another potential explanation for our inability to replicate our observations across cohorts is the fact that our power to detect these effects is reduced due to the small sample sizes used here. Although we have demonstrated that we have 80% power in our discovery cohort to detect large changes in gene expression ( $\Delta = 0.52$ ), this is for our observed 90th percentile of the standard deviation for all genes from our data ( $\sigma = 0.53$ ) (electronic supplementary material, figure S2); probes with smaller standard deviations would not be detectable, either for these probes in replication cohorts or for other probes in the discovery cohort. Thus, we verified our findings in a neuroblastoma cell culture model for two genes, which were predicted to be targets of miR-484, but provided inconsistent observations. The first, *TRIOBP*, was observed in our previously published study of the publicly available data on the CEU (Utah residents with North and Western European ancestry) individuals [30] and was replicated in our discovery family cohort, but not in either the twin or GTEx replication cohorts. In contrast, the second, *NDE1*, was not observed in any of the Finnish cohorts but was strongly implicated in the

larger GTEx data where *NDE1* expression levels were strongly associated with the *NDE1* SNP rs2242549 genotype ( $\beta = 0.24$ ; s.e. = 0.037;  $t$ -statistic = 6.6;  $p = 2.3 \times 10^{-10}$ ). The proteins encoded for by these genes were each found to be significantly increased by the presence of the mimic of miR-484. Such a verification analysis would be required for all genes implicated in this study. However, with such a large number of genes identified this was not feasible with the cell culture model used here.

When the *DISC1* network was studied with respect to treatment, the *NDE1* locus again demonstrated association, specifically in interaction with gender for drugs metabolized by CYP2C19. The degree of expression of cytochrome P450 enzymes in lymphocytes was too low to allow us to investigate potential changes in expression level in our family data. Instead we demonstrated in a cell culture model that miR-484 is capable of increasing the expression of CYP2C19, but not that of another major metabolizing enzyme for psychoactive drugs, CYP2D6. The miRWalk database does not robustly predict either enzyme to be a target for miR-484, thus the mechanism of this action needs to be further investigated. The results from the cell culture assays clearly indicate, however, that miR-484 is capable of affecting the expression level of CYP2C19, suggesting that the mechanism through which the locus confers risk and alters medication usage could be the same. In spite of the lack of confident prediction, a direct effect of miR-484 on the *CYP2C19* locus remains the simplest explanation for these observations, although indirect effects remain possible.

In the case of medication use, we employed a dichotomous variable based on a cut-off of ceasing to use the prescribed medication after three months or less. This was designed to indicate that a treatment was either not considered to be working or else was having side effects which were too severe and its use was, therefore, stopped. Since the cell culture experiment showed that CYP2C19 protein expression is increased by miR-484, it can be hypothesized that the medications are more rapidly metabolized in individuals carrying these variants, leading to a reduced efficacy of those treatments. Interestingly, the genetic effects on medication differ depending on both class of drug and gender. Males carrying the risk alleles had a higher probability of cessation of use for SSRIs metabolized by CYP2C19, while females carrying the risk alleles had an increased probability of cessation for non-SSRI drugs that are metabolized by CYP2C19. Although the mechanism for these effects remains unclear at this time, it is noteworthy that the original association between the *NDE1* locus and schizophrenia in these families was significant only in females [3]. Taken together this implies that one or more gender-specific effects act as modifying factors in conjunction with the *NDE1*/miR-484 locus, although this cannot be easily modelled in our cell culture system.

Here, through the identification of altered gene expression patterns that led to the functional implication of miR-484, which is coded on an untranslated exon of the primary isoform of *NDE1*, we identified a means by which genetic variation in the *DISC1* network can not only increase risk of major mental illnesses, but also how those same variants can alter treatment response to specific psychoactive medications through the regulation of their metabolizing enzyme. This study has, therefore, provided new biological insight into psychiatric disorders for which novel medications could be designed, as well as suggesting that

knowledge of an individual's genotype within the *NDE1*/miR-484 locus may have potential value in the targeting of current therapies.

**Ethics.** In all studies, the principles recommended in the Declaration of Helsinki and its amendments were followed. The study has been approved by the Coordinating Ethics committee of the Hospital District of Helsinki and Uusimaa. Informed consent was obtained from all participants.

**Data accessibility.** The discovery cohort used here is ascertained based on schizophrenia, a disorder that remains highly stigmatizing to the patients and their families. Thus, anonymized gene expression data from this cohort are held by the NCBI Gene Expression Omnibus (accession no. GSE48072). All other details relating to this cohort, which includes information that could be used to identify the patients, are available to any scientist through application to the biobank in which they are stored (<https://www.thl.fi/en/web/thlfi-en/topics/information-packages/thl-biobank>).

**Authors' contributions.** N.J.B., L.U.V. and W.H. wrote the manuscript text; N.J.B., M.P., L.U.V. and W.H. prepared the manuscript figures; N.J.B.,

J.S., J.L., J.H. and W.H. designed the study; T.P., J.S., J.L., T.D.C. and J.H. provided access to samples and data; N.J.B. performed laboratory experiments; N.J.B., L.U.V., M.P., A.B.Z., A.O.A., M.T.H. V.S. and W.H. performed the analysis; all authors reviewed the manuscript.

**Competing interests.** We have no competing interests.

**Funding.** This work has been supported by the Academy of Finland (grant nos.: 128504, 259589, 265097 to W.H.), EU-FP7 (MC-ITN no. 607616 'IN-SENS' to W.H.), the Orion Farnos Research Foundation (to W.H.), the Forschungskommission of the Heinrich Heine University Medical Faculty (grant no.: 9772547 to N.J.B.), the Fritz Thyssen Foundation (grant no.: 10.14.2.140 to N.J.B.) and the Alexander von Humboldt Foundation (fellowship no.: 1142747 to N.J.B.).

**Acknowledgements.** Gene expression analysis was performed by the Institute for Molecular Medicine Finland FIMM Technology Centre, University of Helsinki. Jaakko Kaprio is gratefully acknowledged for the provision of control twins to this study and for critical reading of the manuscript. Antti Tanskanen of the National Institute for Health and Welfare is thanked for producing the definitions of the medication periods used in this study.

## References

- Hennah W *et al.* 2003 Haplotype transmission analysis provides evidence of association for DISC1 to schizophrenia and suggests sex-dependent effects. *Hum. Mol. Genet.* **12**, 3151–3159. (doi:10.1093/hmg/ddg341)
- Ekelund J, Hennah W, Hiekkalinna T, Parker A, Meyer J, Lönnqvist J, Peltonen L. 2004 Replication of 1q42 linkage in Finnish schizophrenia pedigrees. *Mol. Psychiatry* **9**, 1037–1041. (doi:10.1038/sj.mp.4001536)
- Hennah W *et al.* 2007 Families with the risk allele of DISC1 reveal a link between schizophrenia and another component of the same molecular pathway, NDE1. *Hum. Mol. Genet.* **16**, 453–462. (doi:10.1093/hmg/ddl462)
- Tomppo L *et al.* 2009 Association between genes of disrupted in schizophrenia 1 (DISC1) interactors and schizophrenia supports the role of the DISC1 pathway in the etiology of major mental illnesses. *Biol. Psychiatry* **65**, 1055–1062. (doi:10.1016/j.biopsych.2009.01.014)
- Millar JK, Christie S, Porteous DJ. 2003 Yeast two-hybrid screens implicate DISC1 in brain development and function. *Biochem. Biophys. Res. Comm.* **311**, 1019–1025. (doi:10.1016/j.bbrc.2003.10.101)
- Morris JA, Kandpal G, Ma L, Austin CP. 2003 DISC1 (Disrupted-In-Schizophrenia 1) is a centrosome-associated protein that interacts with MAP1A, MIPT3, ATF4/5 and NUDEL: regulation and loss of interaction with mutation. *Hum. Mol. Genet.* **12**, 1591–1608. (doi:10.1093/hmg/ddg162)
- Ozeki Y *et al.* 2003 Disrupted-in-schizophrenia-1 (DISC-1): mutant truncation prevents binding to NudE-like (NUDEL) and inhibits neurite outgrowth. *Proc. Natl Acad. Sci. USA* **100**, 289–294. (doi:10.1073/pnas.0136913100)
- Millar JK *et al.* 2005 DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signalling. *Science* **310**, 1187–1191. (doi:10.1126/science.1112915)
- Palo OM *et al.* 2007 Association of distinct allelic haplotypes of DISC1 with psychotic and bipolar spectrum disorders and with underlying cognitive impairments. *Hum. Mol. Genet.* **16**, 2517–2528. (doi:10.1093/hmg/ddm207)
- Hennah W *et al.* 2009 DISC1 association, heterogeneity and interplay in schizophrenia and bipolar disorder. *Mol. Psychiatry* **14**, 865–873. (doi:10.1038/mp.2008.22)
- Kilpinen H *et al.* 2008 Association of DISC1 with autism and Asperger syndrome. *Mol. Psychiatry* **13**, 187–196. (doi:10.1038/sj.mp.4002031)
- Hennah W, Tuulio-Henriksson A, Paunio T, Ekelund J, Varilo T, Partonen T, Cannon TD, Lönnqvist J, Peltonen L. 2005 A haplotype within the DISC1 gene is associated with visual memory functions in families with a high density of schizophrenia. *Mol. Psychiatry* **10**, 1097–1103. (doi:10.1038/sj.mp.4001731)
- Cannon TD *et al.* 2005 Association of DISC1/TRAX haplotypes with schizophrenia, reduced prefrontal gray matter, and impaired short- and long-term memory. *Arch. Gen. Psychiatry* **62**, 1205–1213. (doi:10.1001/archpsyc.62.11.1205)
- Wegelius A, Pankakoski M, Tomppo L, Lehto U, Lönnqvist J, Suvisaari J, Paunio T, Hennah W. 2015 An interaction between NDE1 and high birth weight increases schizophrenia susceptibility. *Psychiatry Res.* **230**, 194–199. (doi:10.1016/j.psychres.2015.08.038)
- Sullivan PF. 2013 Questions about DISC1 as a genetic risk factor for schizophrenia. *Mol. Psychiatry* **18**, 1050–1052. (doi:10.1038/mp.2012.182)
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014 Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427. (doi:10.1038/nature13595)
- Millar JK *et al.* 2000 Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* **9**, 1415–1423. (doi:10.1093/hmg/9.9.1415)
- Thomson PA *et al.* 2016 Balanced translocation linked to psychiatric disorder, glutamate, and cortical structure/function. *NPJ Schizophr.* **2**, 16024. (doi:10.1038/npjscz.2016.24)
- Sahoo T *et al.* 2011 Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genet. Med.* **13**, 868–880. (doi:10.1097/GIM.0b013e3182217a06)
- Malhotra D, Sebat J. 2012 CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* **148**, 1223–1241. (doi:10.1016/j.cell.2012.02.039)
- Rees E *et al.* 2014 Analysis of copy number variations at 15 schizophrenia-associated loci. *Br. J. Psychiatry* **204**, 108–114. (doi:10.1192/bjp.bp.113.131052)
- Rodríguez-López J *et al.* 2015 An efficient screening method for simultaneous detection of recurrent copy number variants associated with psychiatric disorders. *Clin. Chim. Acta* **445**, 34–40. (doi:10.1016/j.cca.2015.03.013)
- Kimura H *et al.* 2014 Identification of rare, single-nucleotide mutations in NDE1 and their contributions to schizophrenia susceptibility. *Schizophr. Bull.* **41**, 744–753. (doi:10.1093/schbul/sbu147)
- Alkuray F *et al.* 2011 Human mutations in NDE1 cause extreme microcephaly with lissencephaly. *Am. J. Hum. Genet.* **88**, 536–547. (doi:10.1016/j.ajhg.2011.04.003)
- Bakircioglu M *et al.* 2011 The essential role of centrosomal NDE1 in human cerebral cortex neurogenesis. *Am. J. Hum. Genet.* **88**, 523–535. (doi:10.1016/j.ajhg.2011.03.019)
- Güven A, Gunduz A, Bozoglu T, Yalcinkaya C, Tolun A. 2012 Novel NDE1 homozygous mutation resulting in microhydranencephaly and not microlyssencephaly. *Neurogenetics* **13**, 189–194. (doi:10.1007/s10048-012-0326-9)
- Paciorkowski AR *et al.* 2013 Deletion 16p13.11 uncovers NDE1 mutations on the non-deleted

- homolog and extends the spectrum of severe microcephaly to include fetal brain disruption. *Am. J. Med. Genet.* **161A**, 1523–1530. (doi:10.1002/ajmg.a.35969)
28. de Kovel CGF *et al.* 2010 Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* **133**, 23–32. (doi:10.1093/brain/awp262)
  29. Fujitani M, Zhang S, Fujiki R, Fujihara Y, Yamashita T. 2016 A chromosome 16p13.11 microduplication causes hyperactivity through dysregulation of miR-484/protocadherin-19 signaling. *Mol. Psychiatry* **22**, 364–374. (doi:10.1038/mp.2016.106)
  30. Hennah W, Porteous D. 2009 The DISC1 pathway modulates expression of neurodevelopmental, synaptogenic and sensory perception genes. *PLoS ONE* **4**, e4906. (doi:10.1371/journal.pone.0004906)
  31. Hovatta I *et al.* 1998 Linkage analysis of putative schizophrenia gene candidate regions on chromosomes 3p, 5q, 6p, 8p, 20p and 22q in a population-based sampled Finnish family set. *Mol. Psychiatry* **3**, 452–457. (doi:10.1038/sj.mp.4000443)
  32. American Psychiatric Association. 1994 *Diagnostic and statistical manual of mental disorders*, 4th edn. Washington, DC: American Psychiatric Press.
  33. Orešič M *et al.* 2012 Phospholipids and insulin resistance in psychosis: a lipidomics study of twin pairs discordant for schizophrenia. *Genome. Med.* **4**, 1. (doi:10.1186/gm300)
  34. GTEx Consortium. 2015 The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660. (doi:10.1126/science.1262110)
  35. Wedenoja J *et al.* 2008 Replication of linkage on chromosome 7q22 and association of the regional *Reelin* gene with working memory in schizophrenia families. *Mol. Psychiatry* **13**, 673–684. (doi:10.1038/sj.mp.4002047)
  36. Stoll G *et al.* 2013 Deletion of TOP3 $\beta$ , a component of FMRP-containing mRNPs, contributes to neurodevelopmental disorders. *Nat. Neurosci.* **16**, 1228–1237. (doi:10.1038/nn.3484)
  37. Zheutlin AB, Viehman RW, Fortgang R, Borg J, Smith DJ, Suvisaari J, Therman S, Hultman CM, Cannon TD. 2016 Cognitive endophenotypes inform genome-wide expression profiling in schizophrenia. *Neuropsychology* **30**, 40–52. (doi:10.1037/neu0000244)
  38. Bates D, Mächler M, Bolker BM, Walker SC. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48. (doi:10.18637/jss.v067.i01)
  39. Storey JD, Tibshirani R. 2003 Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445. (doi:10.1073/pnas.1530509100)
  40. Liu P, Hwang JTG. 2007 Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. *Bioinformatics* **23**, 739–746. (doi:10.1093/bioinformatics/btl664)
  41. Dweep H, Gretz N. 2015 miRWalk2.0: a comprehensive atlas of microRNA-target interactions. *Nat. Methods* **12**, 697. (doi:10.1038/nmeth.3485)
  42. Haukka J, Suvisaari J, Tuulio-Henriksson A, Lönngqvist J. 2007 High concordance between self-reported medication and official prescription database information. *Eur. J. Clin. Pharmacol.* **63**, 1069–1074. (doi:10.1007/s00228-007-0349-6)
  43. Mantel-Teeuwisse AK, Klungel OH, Verschuren WMM, Porsius A, de Boer A. 2001 Comparison of different methods to estimate prevalence of drug use by using pharmacy records. *J. Clin. Epidemiol.* **54**, 1181–1186. (doi:10.1016/S0895-4356(01)00396-1)
  44. Yan J, Fine J. 2004 Estimating equations for association structures. *Stat. Med.* **23**, 859–874. (doi:10.1002/sim.1650)
  45. R Core Team. 2005 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
  46. Riazuddin S *et al.* 2006 Mutations in *TRIOBP*, which encodes a putative cytoskeletal-organizing protein, are associated with nonsyndromic recessive deafness. *Am. J. Hum. Genet.* **78**, 137–143. (doi:10.1086/499164)
  47. Shahin H *et al.* 2006 Mutations in a novel isoform of *TRIOBP* that encodes a filamentous-actin binding protein are responsible for DFNB28 recessive nonsyndromic hearing loss. *Am. J. Hum. Genet.* **78**, 144–152. (doi:10.1086/499495)
  48. Denson J, Wu Y, Yang W, Zhang J. 2005 Inter-individual variation of several cytochrome P450 2D6 splice variants in human liver. *Biochem. Biophys. Res. Commun.* **330**, 498–504. (doi:10.1016/j.bbrc.2005.03.010)
  49. Bradshaw NJ, Yerabham ASK, Marreiros R, Zhang T, Nagel-Steger L, Korth C. 2017 An unpredicted aggregation-critical region of the actin-polymerizing protein *TRIOBP-1/Tara*, determined by elucidation of its domain structure. *J. Biol. Chem.* **292**, 9583–9598. (doi:10.1074/jbc.M116.767939)
  50. Dietrich M, Grundmann L, Kurr K, Valinotto L, Saussele T, Schmid RD, Lange S. 2005 Recombinant production of human microsomal cytochrome P450 2D6 in the methylotrophic yeast *Pichia pastoris*. *ChemBioChem* **6**, 2014–2022. (doi:10.1002/cbic.200500200)
  51. Bradshaw NJ, Hennah W, Soares DC. 2013 NDE1 and NDEL1: twin neurodevelopmental proteins with similar ‘nature’ but different ‘nurture’. *Biomol. Concepts* **4**, 447–464. (doi:10.1515/bmc-2013-0023)
  52. Barrett JC, Fry B, Maller J, Daly MJ. 2005 Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265. (doi:10.1093/bioinformatics/bth457)
  53. Gabriel SB *et al.* 2002 The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229. (doi:10.1126/science.1069424)
  54. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, Altman RB, Klein TE. 2012 Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **92**, 414–417. (doi:10.1038/clpt.2012.96)
  55. Law V *et al.* 2014 DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* **42**, D1091–D1097. (doi:10.1093/nar/gkt1068)
  56. Hicks JK *et al.* 2015 Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for *CYP2D6* and *CYP2C19* genotypes and dosing of selective serotonin reuptake inhibitors. *Clin. Pharmacol. Ther.* **98**, 127–134. (doi:10.1002/cpt.147)
  57. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017 KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D61. (doi:10.1093/nar/gkw1092)
  58. Bradshaw NJ. 2016 Cloning of the promoter of *NDE1*, a gene implicated in psychiatric and neurodevelopmental disorders through copy number variation. *Neuroscience* **324**, 262–270. (doi:10.1016/j.neuroscience.2016.03.018)
  59. Beveridge NJ, Gardiner E, Carroll AP, Tooney PA, Cairns MJ. 2010 Schizophrenia is associated with an increase in cortical microRNA biogenesis. *Mol. Psychiatry* **15**, 1176–1189. (doi:10.1038/mp.2009.84)
  60. de Rie D *et al.* 2017 An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat. Biotech.* **35**, 872–878. (doi:10.1038/nbt.3947)
  61. Mosca S, Raponi M, Meneghello A, Buratti E, Woods CG, Baralle D. 2017 Human *NDE1* splicing and mammalian brain development. *Sci. Rep.* **7**, 43504. (doi:10.1038/srep43504)
  62. Bradshaw NJ, Hayashi MAF. 2017 NDE1 and NDEL1 from genes to (mal)functions: parallel but distinct roles impacting on neurodevelopmental disorders and psychiatric illness. *Cell. Mol. Life. Sci.* **74**, 1191–1210. (doi:10.1007/s00018-016-2395-7)
  63. Porteous DJ *et al.* 2014 DISC1 as a genetic risk factor for schizophrenia and related major mental illness: response to Sullivan. *Mol. Psychiatry* **19**, 141–143. (doi:10.1038/mp.2013.160)